MULTI-SCALE SEGMENTATION OF RETINAL IMAGES

Artur V. Cideciyan[1,2], Samuel G. Jacobson[1], Joachim H. Nagel[2]

[1]Department of Ophthalmology, Bascom Palmer Eye Institute, School of Medicine
[2]Department of Biomedical Engineering, College of Engineering
University of Miami, Miami, Florida, USA

ABSTRACT

Tiny bright golden patches can be seen on the dark retinal background of some carriers of X-linked retinitis pigmentosa, an incurable blinding disease. We are interested in analyzing quantitatively these unique "tapetal-like reflex" patches in order to increase our understanding of the cellular mechanisms of the disease. In this paper, we describe the multi-scale thresholding method and its application to the segmentation of the tapetal-like reflex in high resolution digital fundus images. Multi-scale thresholding is a local thresholding method that generates results very similar to that of human intuition. Unlike other local thresholding methods, our method successfully ignores small artifacts in dark regions and simultaneously generates high resolution definitions of objects. Other segmentation applications where there are many bright objects on a darker background should profit from the use of multi-scale thresholding.

INTRODUCTION

Ophthalmology is considered second only to radiology in the application of imaging in the medical field [1]. Among the many imaging modalities, photography of the human fundus (the inside, back portion of the eye) is used commonly due to the transparency of the human eye to visible light and due to the non-invasive nature of the procedure. Fundus photographs allow clinicians to document the appearance and monitor the progression of ophthalmic diseases.

We are using digital image processing techniques to study fundus photographs of a unique feature in the retinas of carriers of the disease X-linked retinitis pigmentosa (XLRP). XLRP is an incurable hereditary disease that leads to progressive degeneration of the retina and blindness by the age 40. Various research approaches have been taken to gain a better understanding of XLRP but the condition remains without known cause and cure [2].

Carriers of XLRP (women with one normal X-chromosome and one X-chromosome with the RP gene) can have abnormal golden reflective patches that appear to originate deep in the retina. This scintillating reflection, the tapetal-like reflex, has been known for over 40 years [3] but a systematic investigation of it has not been published. A quantitative analysis of the size and shape distributions of the golden patches that make up the tapetal-like reflex could shed light onto its exact cellular location. In addition, analysis of the changes in reflex properties over time will clarify and quantify the clinical impression that the reflex fades with age. As the reflex is a manifestation unique to the carrier state of XLRP, better understanding of the reflex should lead to better understanding of XLRP.

In this paper, we present a multi-scale thresholding method whereby the tapetal-like reflex is segmented from high resolution digital fundus images of XLRP carriers. Lacking an anatomical explanation, the tapetal-like reflex is defined as golden patches of high reflectance in the retina. Therefore a thresholding methodology is the most straight forward solution to the problem of segmenting the bright tapetal-like reflex from the dark retinal background. In what follows, we show why conventional thresholding is not satisfactory, and we develop multi-scale thresholding that gives superior results.

METHODS

Image Acquisition

Color positive transparencies (Fujichrome 100 RD) are used to record images of the retina obtained with a Fundus camera (Zeiss FF4). Red, green and blue components (8 bits each) of color digital images are obtained by scanning the transparencies with a high resolution slide scanner (Nikon LS3500). A pixel of the resulting digital image corresponds to a 2.5 μm by 2.5 μm area on the retina. The features of interest (i.e. tapetal-like reflex) are golden colored on a reddish/brownish background. As the amount of green added to red determines the yellowness of an image, the information contained in green images is considered sufficient for segmentation. Therefore, the red and blue images are not used in this analysis.

Image Restoration

We previously developed a complete model of the high resolution digital fundus imaging system, made up of four components; the eye, the camera, the film and the scanner [4]. The complete model consists of linear blurring, point nonlinearities and signal dependent noise sources. For mathematical tractability, the complete model was simplified to a model with linear, space-invariant blurring with additive signal-independent noise. A constrained least squares restoration method [5] along with the simplified model is used to simultaneously reduce noise and obtain a better estimate of the underlying retinal reflection function.

Segmentation

Thresholding, in its most general form, refers to the generation of a binary image $g(x,y)$ from a grey level image $f(x,y)$ using a threshold function $t(x,y)$. Specifically,

$$g(x,y) = \begin{cases} 1 \text{ if } f(x,y) > t(x,y) \\ 0 \text{ if } f(x,y) \le t(x,y) \end{cases} \qquad (1)$$

Thus, assuming $f(x,y)$ consists of light objects on a dark background, the thresholded image, $g(x,y)$, will label pixels belonging to objects and background with 1 and 0, respectively. If $t(x,y)$ is a constant independent of spatial coordinates then the threshold is called *global*, otherwise if $t(x,y)$ varies with spatial coordinates then the threshold is called *local*. In most real world images, slowly varying illumination makes it impossible to determine a global threshold that generates an acceptable segmentation. Therefore, local thresholding is used most often where $t(x,y)$ depends on some property in the neighborhood of pixel (x,y).

A simple choice of a neighborhood property to use for local thresholding is the average value. It is defined as the average of pixel intensities within a square neighborhood of size (2m+1) by (2m+1), centered on the pixel of interest; that is

$$\hat{f}_m(x,y) = \frac{1}{(2m+1)^2} \sum_{x-m}^{x+m} \sum_{y-m}^{y+m} f(x,y) \qquad (2)$$

Equation (2) is defined for all x and y that are greater than (m-1) and less than (N-m) for a NxN image.

Setting the threshold function, $t(x,y)$, equal to the average value generates a segmentation where every pixel with an intensity greater than the local average is labeled as an object and all other pixels are labeled as background. With the possible exception of low-noise, high-contrast images, segmentation of most real images with local thresholding (using any neighborhood property) depends on the choice of m, the neighborhood size. With small m, detected object boundaries correspond closely to the actual object boundaries, but at the same time noise and/or artifacts might be mistaken as objects. With large m, noise and/or artifacts are correctly ignored but clusters of small objects might be segmented as a few large objects. Next, we introduce multi-scale thresholding that combines the useful features of small and large neighborhood segmentations.

In multi-scale thresholding the threshold function, $t(x,y)$, is defined as

$$t(x,y) = \max\{ \hat{f}_{m1}(x,y), \hat{f}_{m2}(x,y) \} \tag{3}$$

where

max{} function returns the larger one of its arguments,

$\hat{f}_{m1}(x,y)$ and $\hat{f}_{m2}(x,y)$ are the average values as defined in (2)

m1 and *m2* are two neighborhood sizes with $m2 > m1$.

This definition of multi-scale thresholding is equivalent to the logical 'AND' of $g_1(x,y)$ and $g_2(x,y)$; the locally thresholded images with neighborhood sizes of *m1* and *m2*, respectively. The multi-scale thresholding method, effectively, allows objects only in relatively light areas of the image. Within the allowed areas, best use of the high resolution is made due to local thresholding with small m. The actual sizes of the two neighborhoods used are chosen based on the effective blurring due to the imaging system and the characteristics of the illumination function. We used 6 and 20 for the small and large values of m, respectively.

RESULTS

The green component of a representative fundus image from a carrier of XLRP is shown in Figure 1. The tapetal-like reflex is clearly visible as many bright patches on a dark background. Two squares marked as (A) and (B) are placed on Figure 1 to attract attention to the regions of expected difficulty with segmentation. Specifically, region (A) corresponds to a dark region which does not contain tapetal-like reflex patches. On the other hand, region (B) corresponds to a cluster consisting of many tapetal-like reflex patches close to each other. A successful segmentation should not only ignore small fluctuations in region (A), but also it should be able to separate the individual patches in region (B).

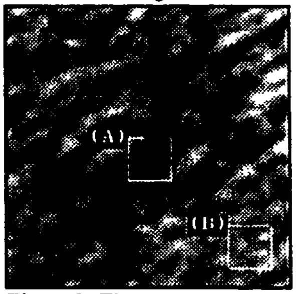

Figure 1: The green component of a 256x256 fundus image showing the tapetal-like reflex.

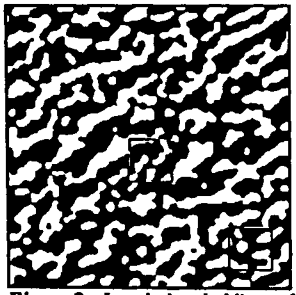

Figure 2: Local thresholding of Figure 1 with m=6.

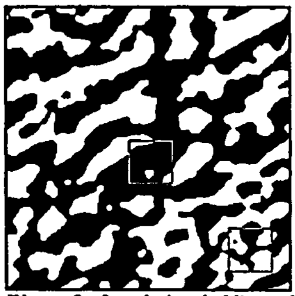

Figure 3: Local thresholding of Figure 1 with m=20.

Results of local thresholding with neighborhood sizes of m=6 and m=20 are shown in Figures 2 and 3, respectively. A look at regions (A) and (B), clearly show the disadvantages of local thresholding with any given fixed neighborhood size; region (A) of Figure 2 is incorrectly filled with objects and the objects in region (B) of Figure 3 are clustered into a pair of large objects. The results of multi-scale thresholding can be seen in Figure 4. The obvious errors apparent in Figures 2 and 3 are drastically reduced and the segmentation corresponds closely to the intuition of an expert human observer (see Discussion). It is important to mention that a morphological 'opening' with a structuring element of size 5x5 was used on all segmentations. Morphological opening achieves the separation of objects sometimes connected with thin 'bridges' [6].

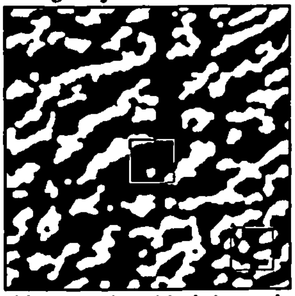

Figure 4: Multi-scale thresholding of Figure 1.

DISCUSSION

The results of the present study indicate that our multi-scale thresholding technique clearly has advantages over more conventional local thresholding methods. Even though we cannot objectively validate the segmentation results, we have made comparisons with the results of manual tracing by an expert. One of us (SGJ), an ophthalmologist specializing in RP, traced the tapetal-like reflex borders on a computer screen with a mouse while referring to a magnified color projection of the fundus photograph. A preliminary comparison of results of manual tracing and our computer segmentation shows quite good agreement (high correlation). Of course, the ease and repeatability of the computer segmentation is a major advantage. Manual segmentation of a single 500x500 pixel subregion took about 10 hours and it was difficult to be certain that the same criteria were applied to all objects.

The future direction of this work will be in the statistical analyses of the patch size and shape distribution of the tapetal-like reflex using a number of retinal regions in XLRP carriers from the same and different families. In addition, the change in the appearance of the reflex will be quantified in XLRP carriers who have been monitored over years (decades, in one case) with serial photographs.

ACKNOWLEDGMENTS

We thank Dr. C.M. Kemp and Dr. R.W. Knighton for critical advice. This research was supported in part by The National Retinitis Pigmentosa Foundation, Inc. (Baltimore, MD) and The Chatlos Foundation, Inc. (Longwood, FL).

REFERENCES

[1] E. Peli, "Electro-optic Fundus Imaging", Surv. Ophthalmol., 34:113-122, 1989.

[2] A.F. Wright, "Towards the Identification of Genes in X-linked Retinitis Pigmentosa". In N. Osborne & G. Chader (eds.), "Progress in Retinal Research", Vol. 9, Oxford, Pergamon Press, 1990, pp.197-227.

[3] H.F. Falls and C.W. Cotterman, "Choroidoretinal Degeneration. A Sex-linked Form in which Heterozygous Women Exhibit a Tapetal-like Retinal Reflex", Arch. Ophthalmol., 40:685-703, 1948.

[4] A.V. Cideciyan, J.H. Nagel, S.G. Jacobson, "Modeling of High Resolution Digital Retinal Imaging", 13th Ann. Intl. Conf. of IEEE/EMBS, 1991 (submitted).

[5] B.R. Hunt, "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer", IEEE Trans. Comp., C-22(9):805-812, 1973.

[6] P. Maragos and R.W. Schafer, "Morphological Systems for Multidimensional Signal Processing", Proc. IEEE, 78(4):690-710, 1990.

Mailing Address: Artur Cideciyan, McKnight Vision Research Center
1638 NW 10th Avenue, Miami, FL 33136, USA
tel: (305)-326-6089, fax: (305)-326-6306

CH3068-4/91/0000-0268 $01.00 © 1991 IEEE